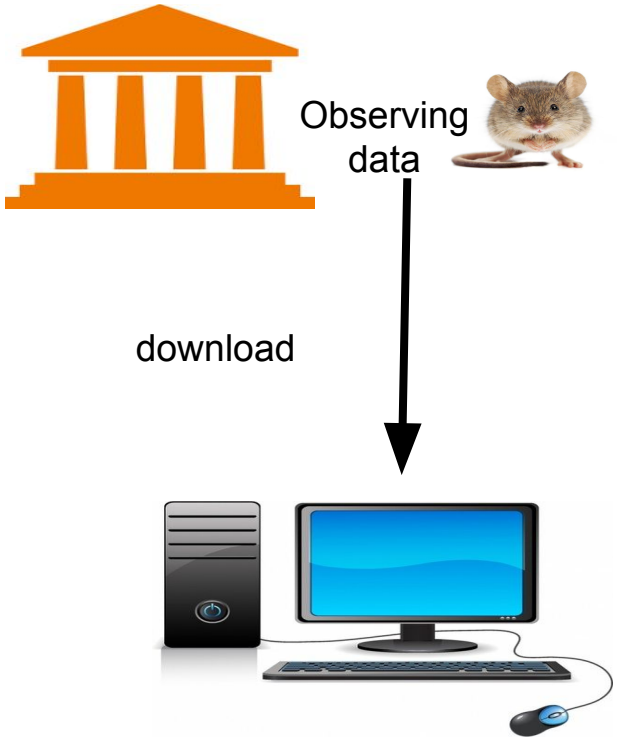
Notebooks examples

Adriano Pieres

Outline

- Why jupyter notebook?
- Examples
- Tips
- Notebooks at NOAO

Why Jupyter Notebooks?



Pros:

- It works for small samples!!!
- The data is close to you!
- You do not have dependence on connectivity, cluster maintenance (?), etc.

Cons:

- Your performance is limited by the facility's connectivity (1), storage (2) and processing feature (3);
- Not shareable; - Crash computer = data loss!

The new deal!

Pros:

- It works for small *and large* samples!!!
- Cluster limitations are set by hardware and by the institution;
- Shareable;
- Code works on the browser;

Cons:

- Dependence of an external institution (maintenance, packages, etc);

A few tips…

https://towardsdatascience.com/the-top-5-magic-commands-for-jupyter-notebooks-2bf0c5ae4bb8

%time, %timeit and %%time

```
import numpy as np
from numpy.random import randint

# A function to simulate one million dice throws.
def one_million_dice():
    return randint(low=1, high=7, size=1000000)

# Let's try %time first
%time throws = one_million_dice()
%time mean = np.mean(throws)

# Outputs:
# Wall time: 20.6 ms
# Wall time: 3.01 ms

# Let's do the same with %timeit
%timeit throws = one_million_dice()
%timeit mean = np.mean(throws)

# Outputs:
# 10 loops, best of 3: 22.2 ms per loop
# 100 loops, best of 3: 2.86 ms per loop

# And finally %%time
%%time
throws = one_million_dice()
mean = np.mean(throws)

# Outputs:
# Wall time: 36.6 ms
```

%matplotlib inline %matplotlib notebook

```
from numpy.random import randint
import matplotlib.pyplot as plt

# Sample 1000 random values to create a scatterplot
x = randint(low=1, high=1000, size=100)
y = randint(low=1, high=1000, size=100)

# This will show nothing in a Jupyter Notebook
plt.scatter(x, y)
plt.show()

# Let the magic happen!
%matplotlib inline
plt.scatter(x, y)
plt.show()
```

%load_ext autoreload

```
# load the autoreload extension
%load_ext autoreload
# Set extension to reload modules every time before executing code
%autoreload 2

from helperfunctions import complicated_function_to_return_a_number

complicated_function_to_return_a_number()
# Output: 123

# Go to helperfunctions.py and change something withing the function

complicated_function_to_return_a_number()
# Output: 321
```

%system

```
# Easy to read version
%system date

# Shorthand with "!!" instead of "%system" works equally well
!!date
```

More tips?

Tips and shortcuts:

Magic Python Documentation

Run Jupyter notebooks for other languages (**Ju**lia **Py**thon and **R**)

```
from IPython.core.interactiveshell import InteractiveShell
InteractiveShell.ast_node_interactivity = "all"
```

```
from pydataset import data
quakes = data('quakes')
quakes.head()
quakes.tail()
```

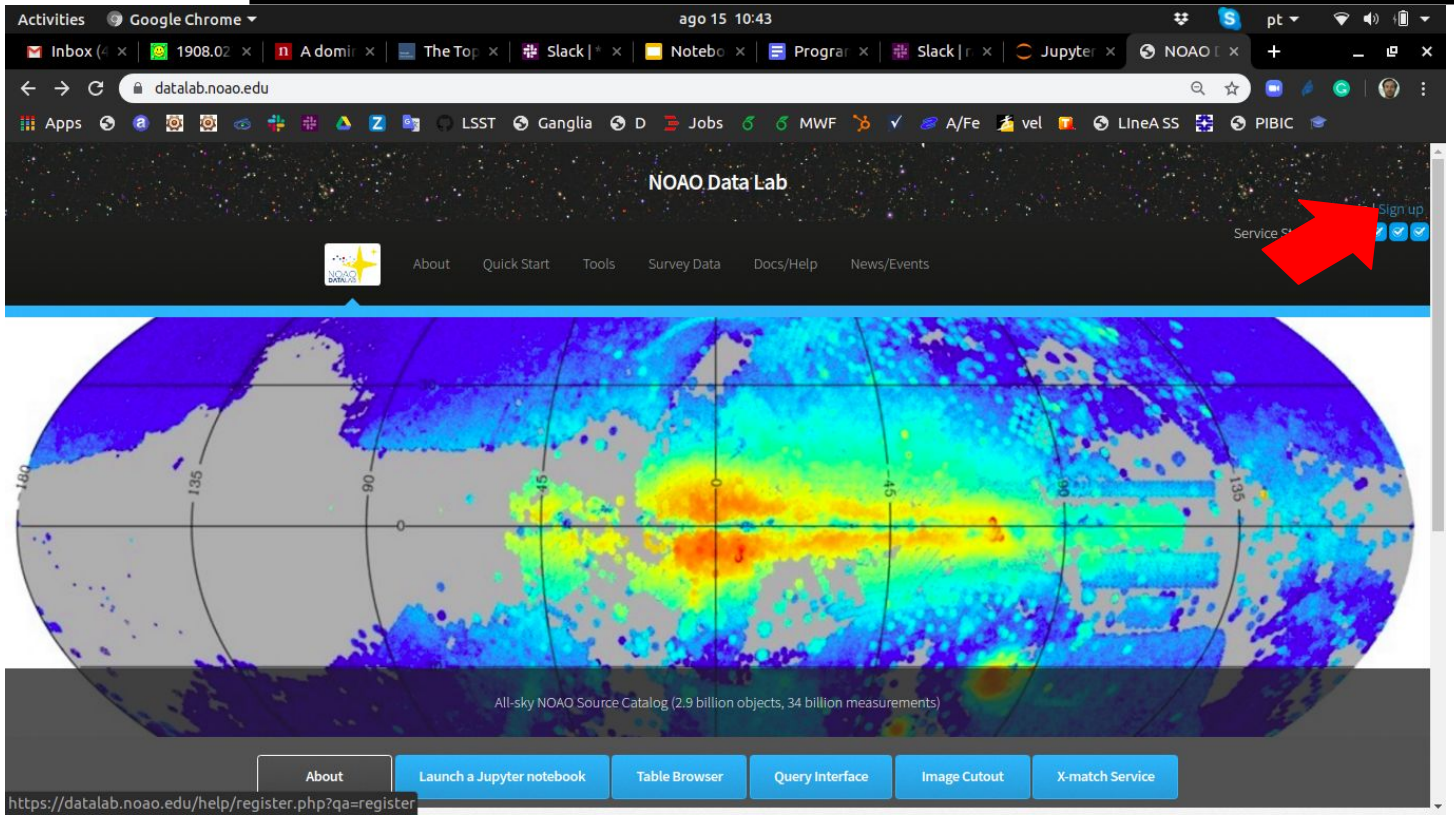
	lat	long	depth	mag	stations
1	-20.42	181.62	562	4.8	41
2	-20.62	181.03	650	4.2	15
3	-26.00	184.10	42	5.4	43
4	-17.97	181.66	626	4.1	19
5	-20.42	181.96	649	4.0	11

	lat	long	depth	mag	stations
996	-25.93	179.54	470	4.4	22

NOAO jupyter notebook

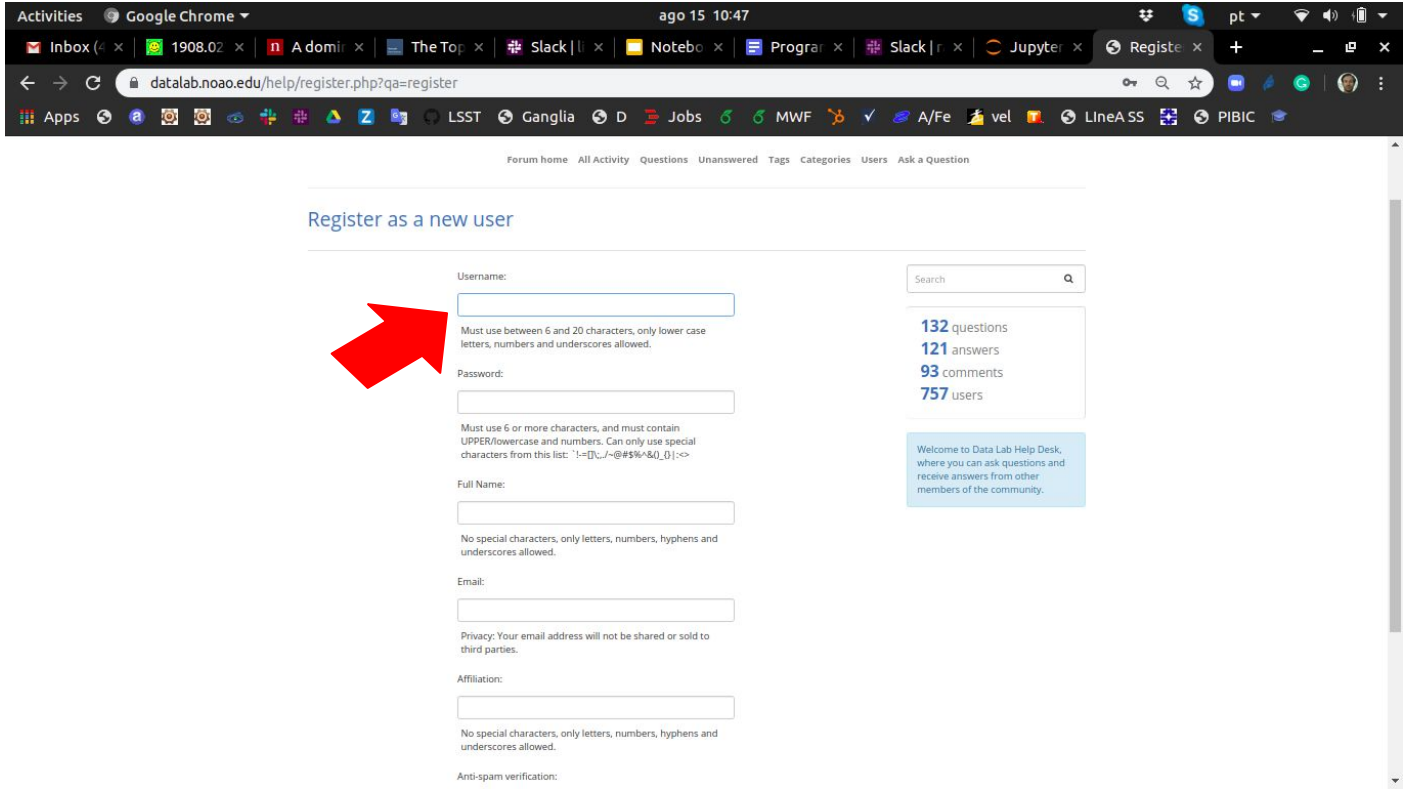
datalab.noao.edu/

Firstly, register yourself!

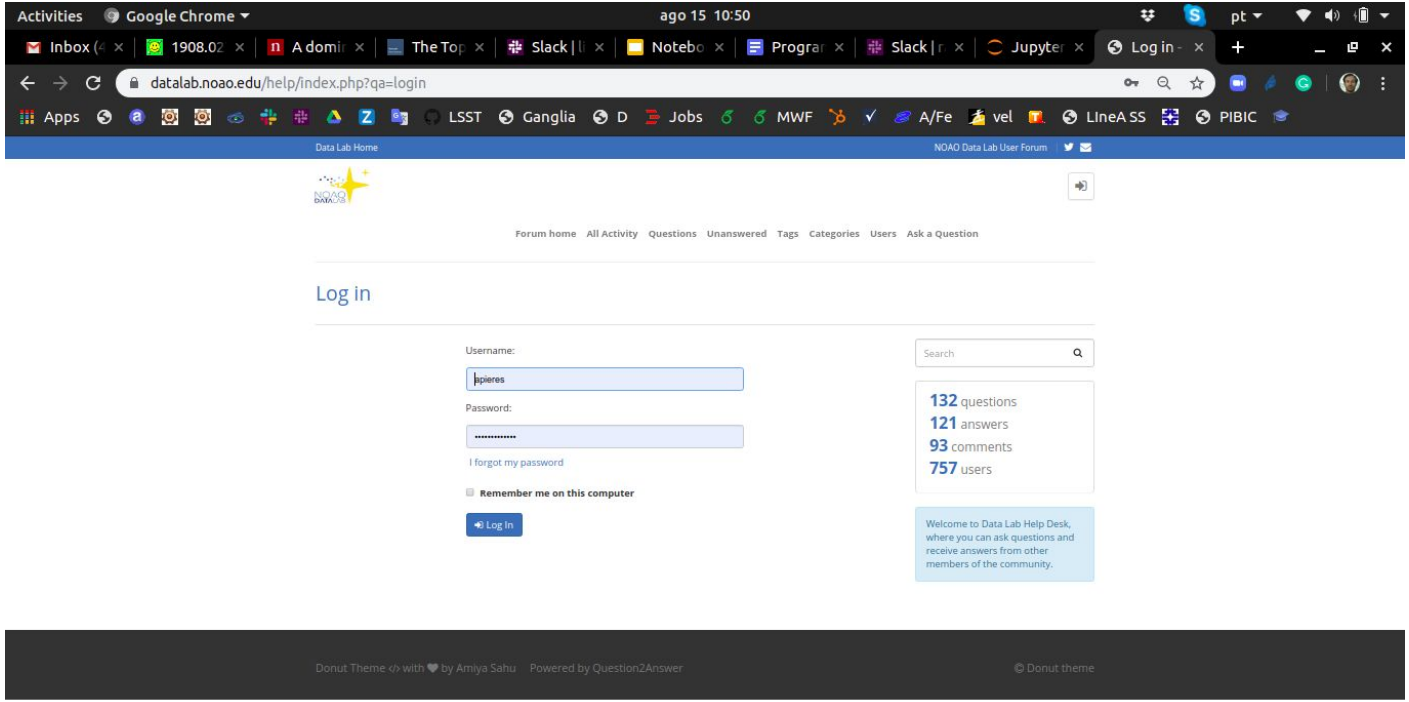


IF I BUILD IT, THEY WILL
COME

Fill the empty fields… you will receive and email...



… and *voilà*: you are registered!



Datalab jupyter notebooks

- Any doubt about registration: quick start, FAQ.
- Data available on datalab:

DES

Legacy Survey

NOAO Source Catalog

SMASH

DECam Asteroid Database

DES SVA1

GAIA

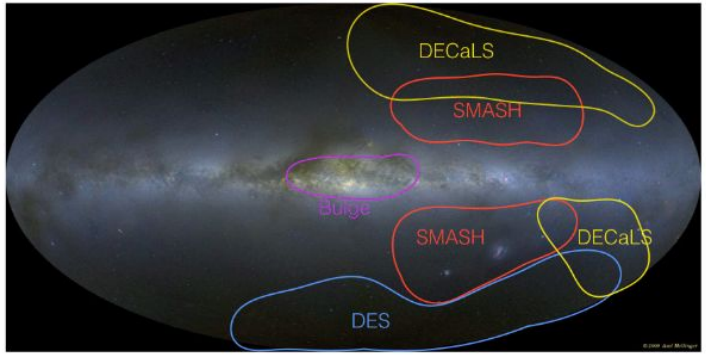
PHAT

SDSS DR13

SPLUS

USNO

The map below shows the areas covered by surveys with catalog data currently available through the NOAO Data Lab (SMASH, DECaLS) and those expected within approximately a year (DES, Bulge). Hover over an outline to see the survey name or click on an outline or on the sidebar links to go to the page for that survey. For pure image data from these and other observing programs, visit the NOAO Science Archive.



NOAO data

DES

Legacy Survey

NOAO Source Catalog

SMASH

DECam Asteroid Database

DES SVA1

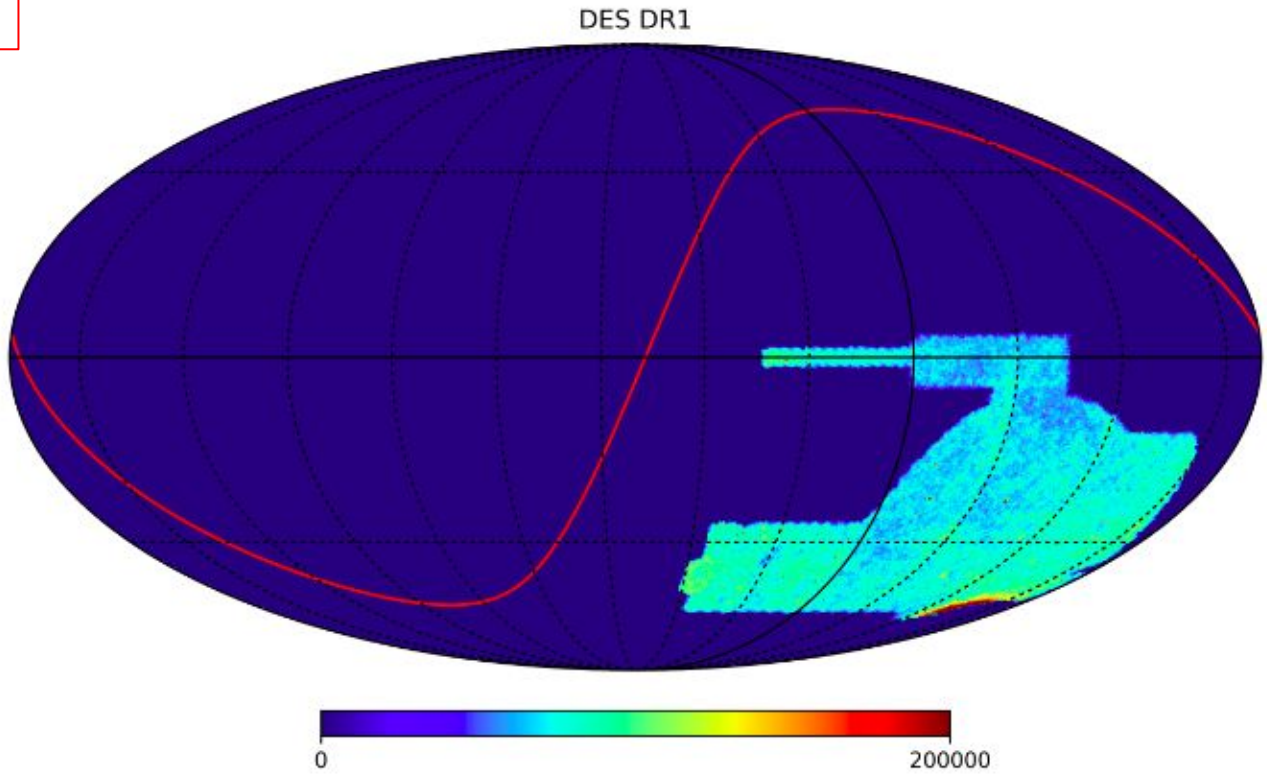
GAIA

PHAT

SDSS DR13

SPLUS

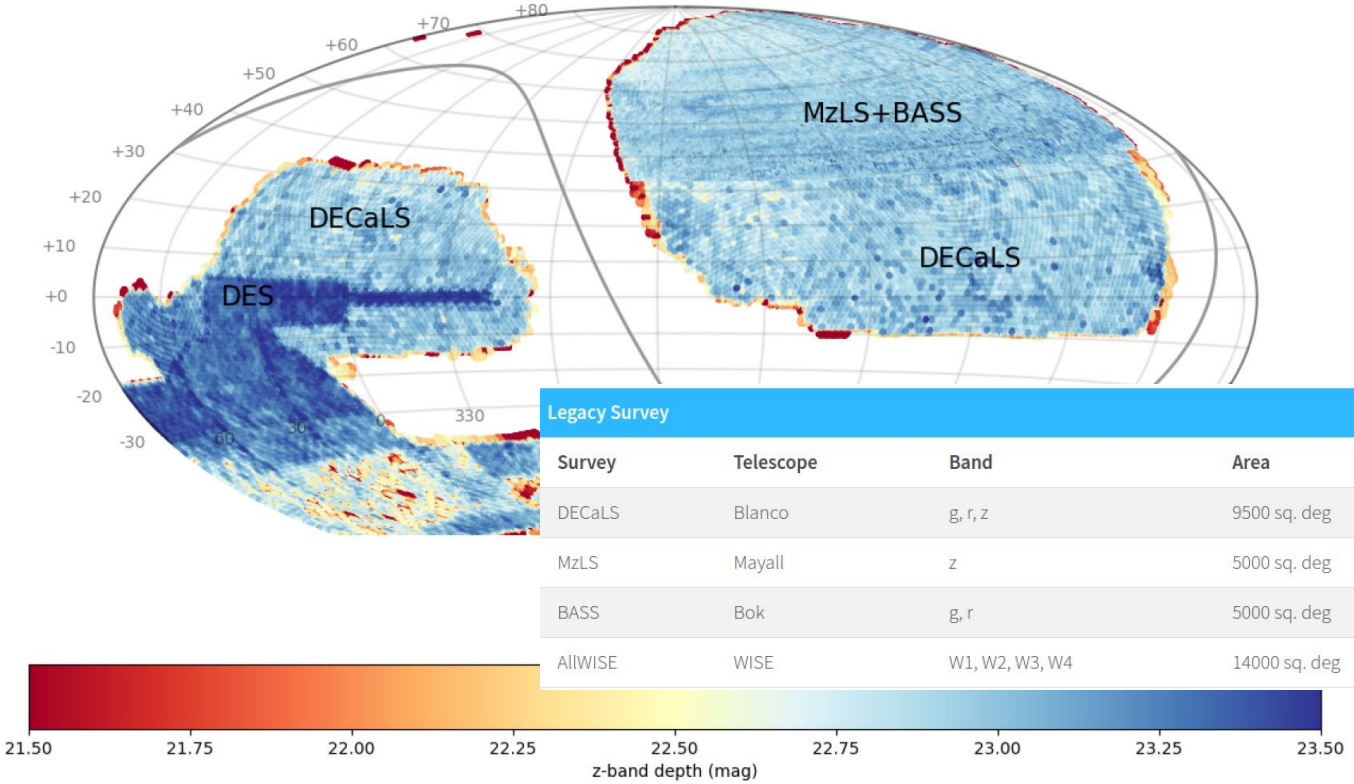
USNO



NOAO data

- DES
- Legacy Survey
- NOAO Source Catalog
- SMASH

- DECam Asteroid Database
- DES SVA1
- GAIA
- PHAT
- SDSS DR13
- SPLUS
- USNO



Legacy Survey

Survey	Telescope	Band	Area
DECaLS	Blanco	g, r, z	9500 sq. deg
MzLS	Mayall	z	5000 sq. deg
BASS	Bok	g, r	5000 sq. deg
AllWISE	WISE	W1, W2, W3, W4	14000 sq. deg

NOAO data

DES

Legacy Survey

NOAO Source Catalog

SMASH

DECam Asteroid Database

DES SVA1

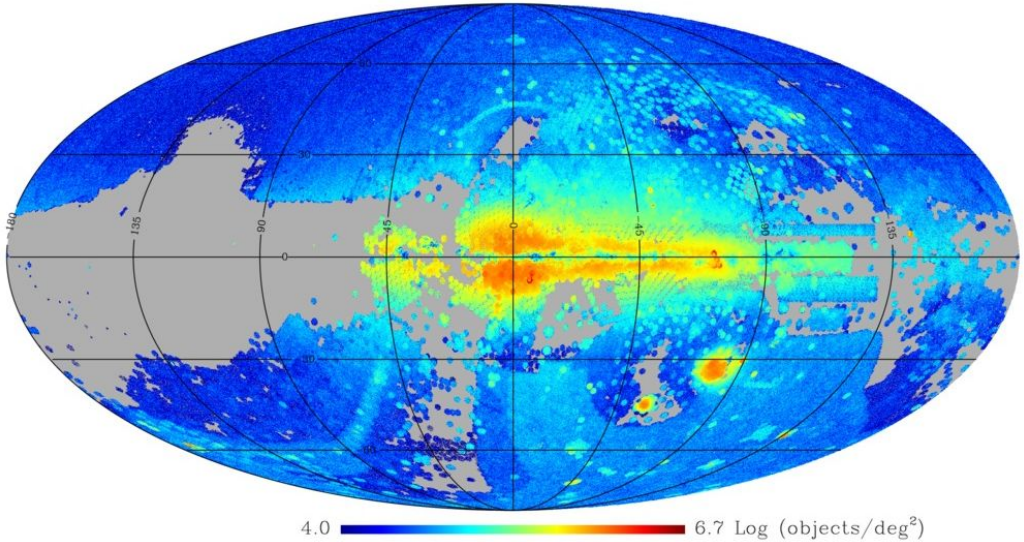
GAIA

PHAT

SDSS DR13

SPLUS

USNO



- Depth: 22.3, 22.8, 22.6, 22.4, 21.7, 20.4 mag (ugrizY)
- Photometric calibration accuracy: ~2%
- Astrometric accuracy: ~20 milliarcsec
- Number of exposures: ~200,000
- Median number of exposures per area: ~15
- Data from DECam, Mosaic-3 and 90Prime

- Source Extractor aperture photometry
- Calibrated using Gaia/2MASS/APASS/Galex
- Object catalog contains:
  - average photometry and uncertainties
  - photometric variability
  - morphology parameters
  - proper motions

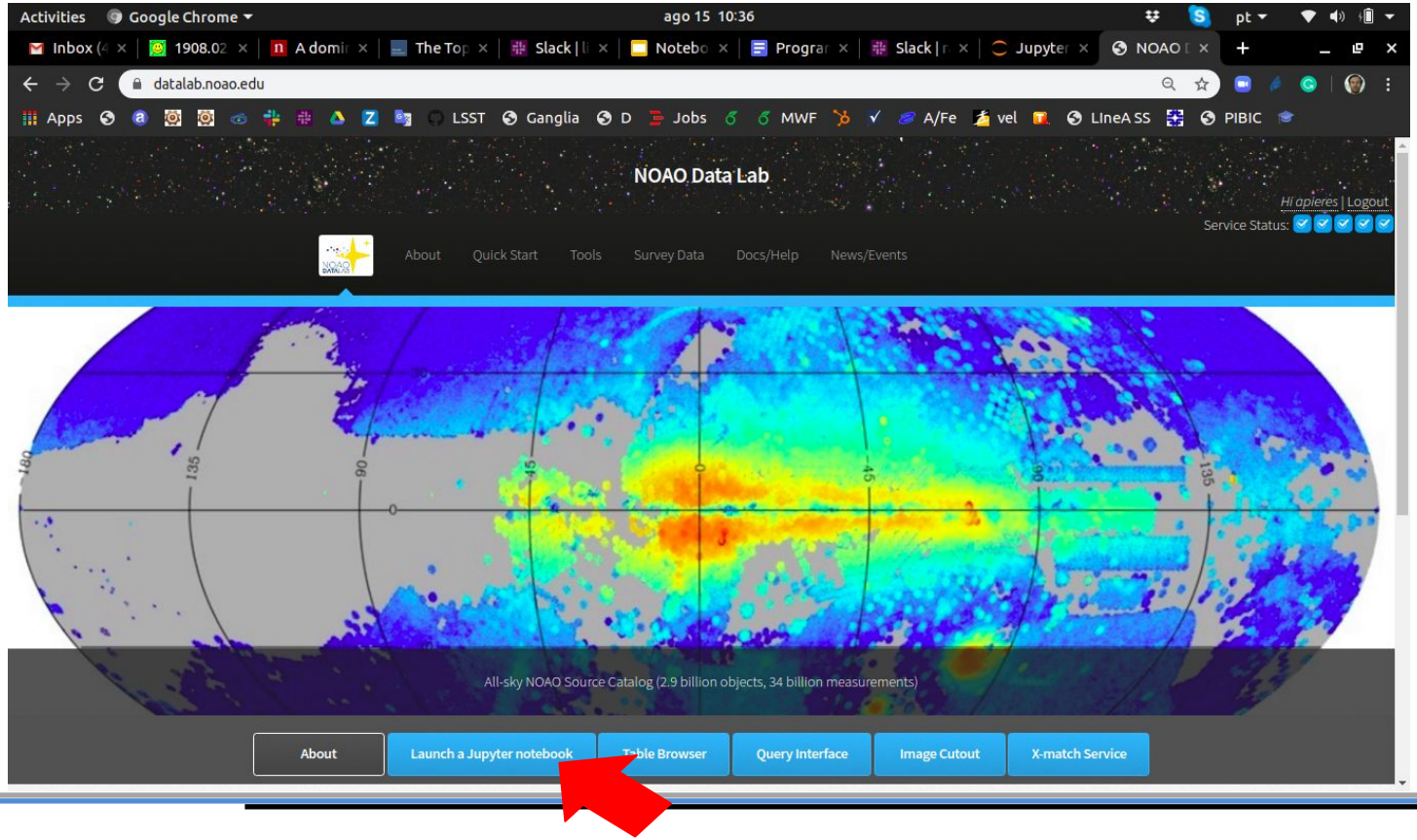
NOAO data

- DES
- Legacy Survey
- NOAO Source Catalog
- SMASH

- DECam Asteroid Database
- DES SVA1
- GAIA
- PHAT
- SDSS DR13
- SPLUS
- USNO

SMASH at a Glance	
Area covered	480 deg^2 spanning ~2400 deg^2
Bands	*ugriz*
Depth (5σ, *ugriz*)	23.9, 24.8, 24.5, 24.2, 23.5 mag
Seeing (*ugriz*)	1.22, 1.13, 1.01, 0.95, 0.90 arcsec
Number of fields	197
- DR1	61
Number of DECam exposures	5,809
- DR1	2,480
Number of objects	~420,000,000
- DR1	101,425,210

Acessing datalab jupyter notebooks (datalab.noao.edu)



● The latest JN available

Go to:
> New > Terminal
Type in terminal:
> getlatest
And you have the latest JN available!

● Latest notebooks available

jupyter

Logout | Control Panel

Files | Running | Clusters

Select items to perform actions on them.

Upload | New

0 /	Name	Last Modified	File size
notebooks		a year ago	
notebooks-latest		a year ago	
notebooks_20190826_083753		6 days ago	
vospace		a year ago	

Latest notebooks available

0 / notebooks_20190826_083753

- ..
- 01_GettingStartedWithDataLab
- 02_DataAccessOverview
- 03_ScienceExamples
- 04_HowTos
- 05_Contrib
- 06_EPO
- DataLabNotebookTemplate.ipynb
- CONTRIBUTING
- keywords.txt
- LICENSE
- README.txt

Do not skip this!!! Datalab packages, SQL...

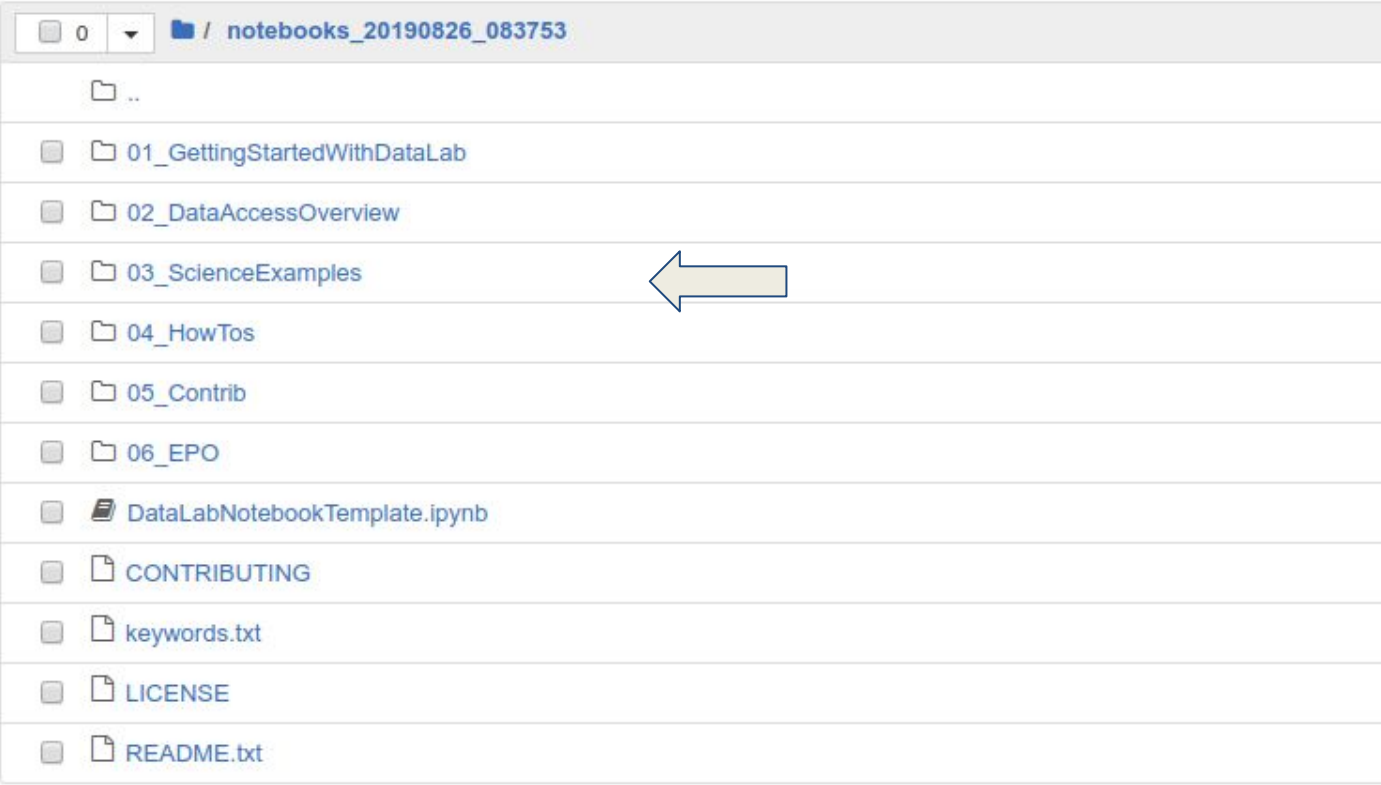
Autentication for datalab, image cutout

● Latest notebooks available

0 / notebooks_20190826_083753

- ..
- 01_GettingStartedWithDataLab
- 02_DataAccessOverview ← Authentication, WISE and SMASH CMDs, light curves
- 03_ScienceExamples
- 04_HowTos
- 05_Contrib
- 06_EPO
- DataLabNotebookTemplate.ipynb
- CONTRIBUTING
- keywords.txt
- LICENSE
- README.txt

● Latest notebooks available



Latest notebooks available

0 / notebooks_20190826_083753 / 03_ScienceExamples

..

DwarfGalaxies

ExploringM31

GalacticStructure

LargeScaleStructure

Pal5TidalTails

SpectralEnergyDistributions

StarGalQSOSeparation

TimeSeriesAnalysisRrLyraeStar

Analyzing DG in DESDR1, Noao Source Catalog, and find in SMASH!

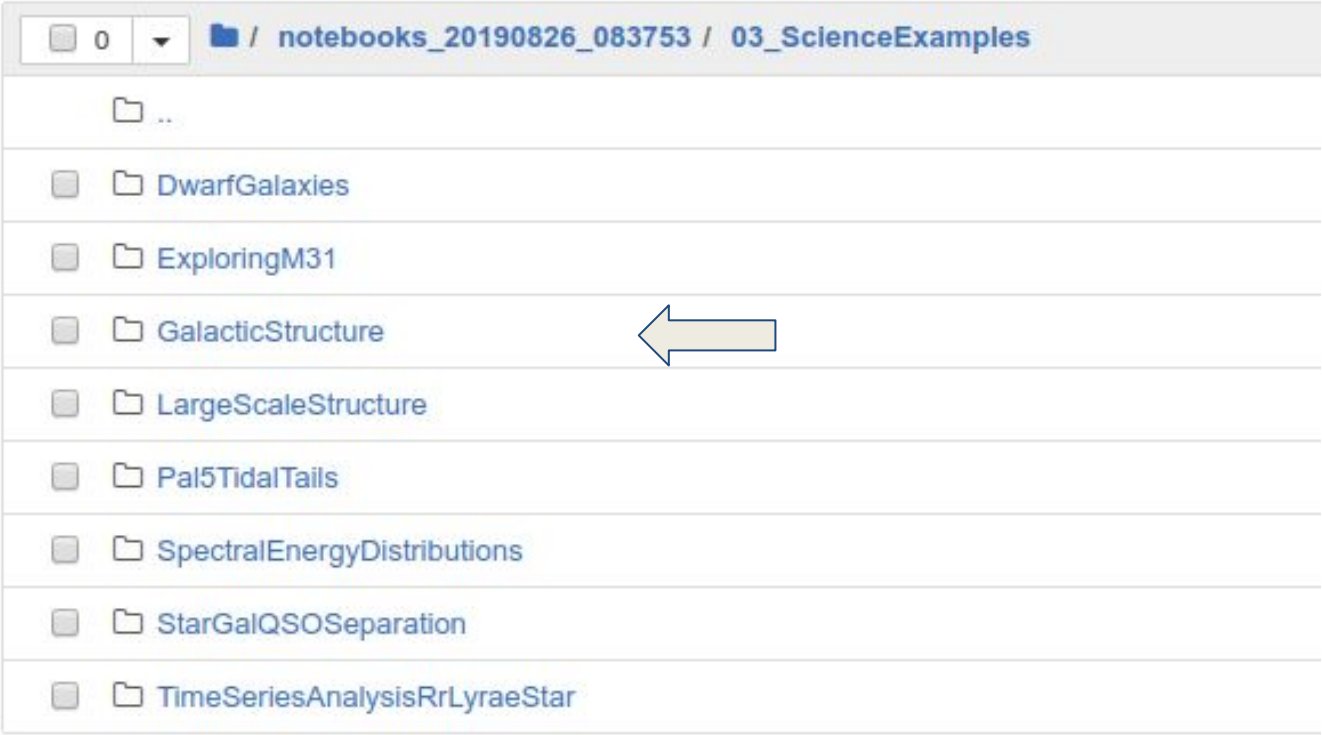
Latest notebooks available

0 / notebooks_20190826_083753 / 03_ScienceExamples

- ..
- DwarfGalaxies
- ExploringM31
- GalacticStructure
- LargeScaleStructure
- Pal5TidalTails
- SpectralEnergyDistributions
- StarGalQSOSeparation
- TimeSeriesAnalysisRrLyraeStar

M31 using Panchromatic Hubble Andromeda Treasury, ⅓ of star forming disk in 6 bands, HP maps of color, CMDs, time series.

● Latest notebooks available



Latest notebooks available

0 / notebooks_20190826_083753 / 03_ScienceExamples / GalacticStructure

..

GalacticStructureDECaPS.ipynb

grizY, 2G stars, 1000 sq deg, SQL, CMDs, CCDs, regions

StarClustersGaiaDR2.ipynb

Dusty Region.png

Field Region.png

Open Cluster.png

StarClustersGaiaDR2.html

Latest notebooks available

0 / notebooks_20190826_083753 / 03_ScienceExamples / GalacticStructure

..

GalacticStructureDECaPS.ipynb

StarClustersGaiaDR2.ipynb

query by SC name, animation plots for NGC6281 and NGC6397 paralaxes

Dusty Region.png

Field Region.png

Open Cluster.png

StarClustersGaiaDR2.html

Latest notebooks available

0 / notebooks_20190826_083753 / 03_ScienceExamples

- ..
- DwarfGalaxies
- ExploringM31
- GalacticStructure
- LargeScaleStructure
- Pal5TidalTails
- SpectralEnergyDistributions
- StarGalQSOSeparation
- TimeSeriesAnalysisRrLyraeStar

DECaLS and SDSS/BOSS, joint query, sanity check, exercise

Latest notebooks available

0 / notebooks_20190826_083753 / 03_ScienceExamples

- ..
- DwarfGalaxies
- ExploringM31
- GalacticStructure
- LargeScaleStructure
- Pal5TidalTails
- SpectralEnergyDistributions
- StarGalQSOSeparation
- TimeSeriesAnalysisRrLyraeStar

Joint query (NSC DR1 and Gaia DR2), proper motion

● Latest notebooks available

0 / notebooks_20190826_083753 / 03_ScienceExamples

- ..
- DwarfGalaxies
- ExploringM31
- GalacticStructure
- LargeScaleStructure
- Pal5TidalTails
- SpectralEnergyDistributions ← Stellar SEDs, CMDs, CCDs with S-PLUS
- StarGalQSOSeparation
- TimeSeriesAnalysisRrLyraeStar

Latest notebooks available

0 / notebooks_20190826_083753 / 03_ScienceExamples

- ..
- DwarfGalaxies
- ExploringM31
- GalacticStructure
- LargeScaleStructure
- Pal5TidalTails
- SpectralEnergyDistributions
- StarGalQSOSeparation
- TimeSeriesAnalysisRrLyraeStar

SG separation in DES DR1 and LS DR7, joint photometry (DECam+WISE) +

Latest notebooks available

0 / notebooks_20190826_083753 / 03_ScienceExamples

- ..
- DwarfGalaxies
- ExploringM31
- GalacticStructure
- LargeScaleStructure
- Pal5TidalTails
- SpectralEnergyDistributions
- StarGalQSOSeparation
- TimeSeriesAnalysisRrLyraeStar

LombScargle periodogram, phased lightcurves

Latest notebooks available

0 / notebooks_20190826_083753 / 04_HowTos

..

AuthClient

CrossmatchTables

FileService

QueryClient

RowVsCstore

SiaService

StoreClient

Latest notebooks available

0 / notebooks_20190826_083753 / 04_HowTos

..

AuthClient

CrossmatchTables

FileService

QueryClient

RowVsCstore

SiaService

StoreClient

0 / notebooks_20190826_083753 / 05_Contrib

..

ExtraGalactic

Galactic

SolarSystem

TimeDomain

Latest notebooks available

0 / notebooks_20190826_083753 / 06_EPO / TeenAstronomyCafe

- ..
- 01_Lives_and_Deaths_of_Stars
- 04_Galactic_Neighborhood
- 05_Gravitational_Lensing
- 06_Black_Holes
- 07_Large_Scale_Structure_of_the_Universe
- README.md

More information about JN datalab available? See README.txt

Suggestion to run

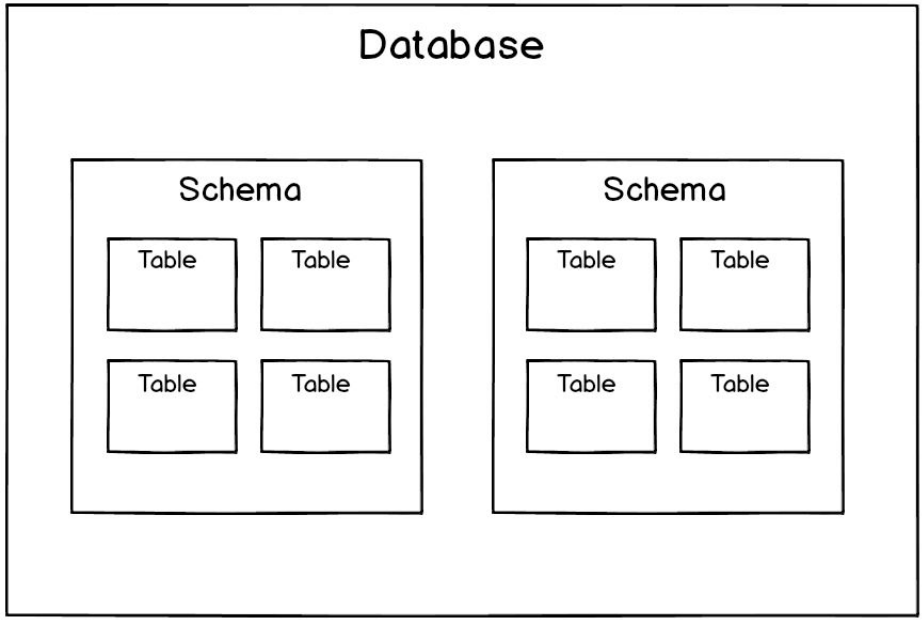
Notebooks > ScienceExamples > StarGalQSOSeparation

References

- NOAO datalab:
  - Fitzpatrick et al., "The NOAO Data Laboratory: a conceptual overview", SPIE, 9149, 2014
  - https://kicp-workshops.uchicago.edu/decam-nfc2018/depot/talk-nikutta-robert.pdf
  - Great opportunity to contribute!

- Jupyter notebooks:
  - jupyter.org
  - https://www.dataquest.io/blog/jupyter-notebook-tips-tricks-shortcuts/

Database, schema, table, column



gaia_source

gaia_dr2

ra, dec, rv, ...

Questões

Obrigado